**Core Argument R Annotation Tool based on Shiny**

**(aka Shiny CARAT)**

by N. Levshina

Version 2 (August 2018)

**User’s guide**

1. Prepare your corpus. Each sentence should be on a separate line. Save it as in the UTF-8 encoding without BOM. You can use a free text editor for this purpose, e.g. Notepad ++.

2. Install RStudio <https://www.rstudio.com/products/rstudio/download/> . A FREE version is perfectly fine.

Version 1 (no longer valid):

Open your corpus in R by using the following code:

yourcorpus <- scan(file = file.choose(), sep = "\n", what= "character", encoding = "UTF-8")

You can check if everything is correct by looking at the first 6 sentences:

head(yourcorpus)

To use the sample corpora, load the .Rdata files in the CARAT directory by clicking on them and then agreeing to load them.

Version 2:

You can use the UI to upload a corpus. There is no need to open it in R first.

3. Install the add-on package *shiny*. To do that, go to the Packages tab, click on Install and type *shiny* in the field “Packages (separate multiple with space or comma):”. The package should be installed automatically. Agree to all default suggestions, if there any. The option “Install dependencies” should be checked.

4. If you work with non-ascii characters, it may be useful to change the locale in R. For example, for Russian, you can write in the R console the following code:

Sys.setlocale(locale = "Russian")

Otherwise, the characters may be represented incorrectly.

5. Copy the directory CARAT from github.com/levshina/CARAT to some directory on your computer.

Version 1 (no longer valid):

Use the “Files” tab to find the directory and click on the file app.R. In the file, add the information about your corpora using . Note: this is a temporary solution. In the future, I want to give the user an opportunity to open corpora directly from the applicaiton.

6. Start the application by pressing the button Run App in the top left panel. You can now code your arguments.

7. Tab “Choose the data”. Here you can choose the corpus and write the name of the file where you want to save the results. Note that everything you save will not be overwritten, but appended to the output file. Use the buttons “Previous” and “Next” to choose the relevant (transitive) sentences. Once you have found a good sentence, press on “Code the sentence.” You will go to the tab “Check predicate”.

8. Tab “Check predicate”. Here you can specify the predicate and check if it meets the criteria. If that’s the case, choose “Yes!” You will go to the tab “Code A”. If not, you can go to the previous tab and choose the next sentence for coding.

9. Tab “Code A”. Fill in the information about A using the lists of features and radio buttons. After you have finish your coding, press “Next”. You will go to the tab “Code P”.

10. Tab “Code P”. The same actions as in “Code A”. After you have finished, press “Next”. You will go to the final tab “Code clause properties”.

11. Tab “Code clause properties.” After you have done your coding, press “Save the results” (a notification should appear) and go to the first tab to choose the next sentence.

12. If you want to finish your work and want to use the same corpora later, click on the cross button in the top right corner and choose “Save”.

Notes: if you have two or more argument structures in one sentence, you can enter them without changing the sentence. Please remember to save your coding each time. Only one pair of core arguments is allowed for one predicate. In case of homogeneous parts (e.g. Mary and John), only the first NP is taken.

Also note that your coding will be appended, not overwritten. This can be changed in the future.